

ChatGPT's Ability to Assist with Clinical Documentation: A Randomized Controlled Trial

Hayden Patrick Baker, Emma Patricia Dwyer, Kelly Hynes, Jennifer Moriatis Wolf¹, Jason Strelzow

¹The University of Chicago

INTRODUCTION:

Clinical documentation is a critical aspect of health care that enables healthcare providers to communicate effectively with each other and maintain accurate patient care records. Artificial intelligence (AI) tools, such as chatbots and virtual assistants, have the potential to assist healthcare providers in clinical documentation. ChatGPT is an AI conversational model that generates human-like responses to text-based prompts. In this study, we sought to investigate ChatGPT's ability to assist with writing a history of present illness (HPI) based on standardized patient histories.

METHODS:

A blinded, randomized controlled study was conducted to compare the use of typing, dictation, and ChatGPT as tools to document HPIs of standardized patient histories. Eleven study participants, consisting of medical students, orthopaedic surgery residents, and attending surgeons, completed 3 HPIs using a different documentation technique for each one. Participants were randomized into cohorts based on the type of documentation technique. Participants were asked to interview standardized patients and document the patient's HPI using their assigned method.

RESULTS:

ChatGPT was found to be intermediate in terms of speed, dictation was fastest, but produced significantly longer and higher-quality patient histories compared to dictation and typing. However, ChatGPT included erroneous information in 36% of the documents. There was poor agreement on the quality of the patient histories between reviewers.

DISCUSSION AND CONCLUSION:

Our study suggests that ChatGPT has the potential to improve clinical documentation by producing more comprehensive and organized HPIs. ChatGPT can generate longer and more detailed documentation compared to traditional documentation methods. However, further studies are needed to investigate and address concerns regarding privacy, bias, and information.

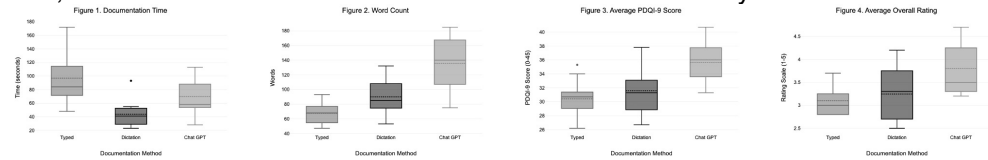


Table 1. Descriptive Statistics

	Typing	Dictation	Chat GPT
Time (sec)	96.8 (39.3)	63.2 (20.2)	69.8 (26.2)
Word Count	67.5 (15.4)	89.9 (21.8)	135.8 (40.3)
POG-I	36.4 (2.7)	31.4 (1.4)	35.8 (2.1)
Overall Rating	3.1 (0.3)	3.3 (0.4)	3.8 (0.4)
History Items	4.9 (0.8)	5 (1.1)	6 (1.1)